

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Fourier Transform Infrared Microspectroscopy as a Tool for Embryonic Stem Cell Studies

Diletta Ami<sup>1,2</sup>, Paolo Mereghetti<sup>3</sup>, Antonino Natalello<sup>1</sup> and Silvia M. Doglia<sup>1</sup>

<sup>1</sup>*Department of Biotechnology and Biosciences, Univeristy of Milano Bicocca*

<sup>2</sup>*Fondazione IRCCS Policlinico San Matteo, Pavia*

<sup>3</sup>*Center for Modelling and Simulation in the Biosciences, University of Heidelberg and HITS gGmbH, Heidelberg*

<sup>1,2</sup>*Italy*

<sup>3</sup>*Germany*

## 1. Introduction

Embryonic stem (ES) cells are self-renewing and pluripotent cells that arise from the inner cell mass of the mammalian blastocyst (Smith, 2001). Their unique capability of potentially generating every cell type continuously attracts the interest of different fields of research. Indeed, ES cells represent a powerful tool for the study of the molecular mechanisms of cell differentiation with important applications in cell therapies, tissue engineering, regenerative medicine and pharmaceutical screening (Trounson, 2006; Vats et al., 2005). All these applications require rapid and sensitive assays to evaluate the differentiation process through the identification of specific markers of the differentiation status. To date, ES cell differentiation is mainly monitored by biochemical methods such as immunohistochemistry, gene expression analysis, functional assays of the differentiating cells and flow cytometry, that - even if providing a comprehensive characterization of cells - are time consuming, expensive and often require a complex sample handling. For these reasons, the development of new approaches for stem cell studies is highly desirable.

In the last decades, optical spectroscopy approaches were applied to the study of intact cells and in particular vibrational spectroscopies revealed to be powerful techniques for the characterization of complex biological systems (Heraud & Tobin, 2009). In particular, Fourier transform infrared (FTIR) and Raman are non invasive and label-free vibrational (micro)spectroscopies that allow to obtain information on the molecular composition and structure of intact cells, tissues and whole organisms (Schulze et al., 2010 ; Tanthanuch et al., 2010; Chan & Lieu, 2009; Walsh et al., 2009; Ami et al., 2004; Choo et al., 1996), providing a unique molecular fingerprint within a single measurement. In this way, it is possible to characterize rapidly different processes that take place simultaneously in biological systems, a non easy task for the standard biochemical approaches. Thanks to the use of an infrared microscope coupled to a FTIR spectrometer, it becomes possible to collect the absorption spectrum from a selected sample area. Interestingly, these techniques, thank to their fast - time resolution, have been successfully used to snapshot and “freeze” molecular events in complex systems (Miller & Dumas, 2010; Hamm, 2009).

The potential of FTIR and Raman spectroscopies has been widely exploited to monitor biological processes in-situ, for instance in cancer diagnosis (Krafft et al., 2009; Baker et al., 2008; Wang et al., 2007), in protein aggregation (Diomedea et al., 2010; Doglia et al., 2008; Ami et al., 2005; Choo et al., 1996) and in stem cell research (Schulze et al., 2010; Heraud & Tobin, 2009; Ami et al., 2008; Notingher et al., 2004a; Notingher et al., 2004b). Indeed, these vibrational approaches have been shown to be a promising tool for the characterization of cellular mechanisms, providing not only structural information, but also details on the dynamics of the structures (Miller & Dumas, 2010).

Concerning in particular the study of complex biological systems, it is important to underline that the multivariate statistical analysis is an essential support to fully understand the spectroscopic response. Among the different statistical approaches, the combined principal component - linear discriminant analysis (PCA-LDA) allows to find in the spectrum the wavenumbers that contribute to the largest inter-spectral variance, thus validating the identification of the marker bands obtained by the direct inspection of the spectral data (Ami et al., 2008; Walsh et al., 2007; Fearn, 2002). In Figure 1 we illustrated the procedure that should be followed to successfully tackle the FTIR characterization of complex systems.

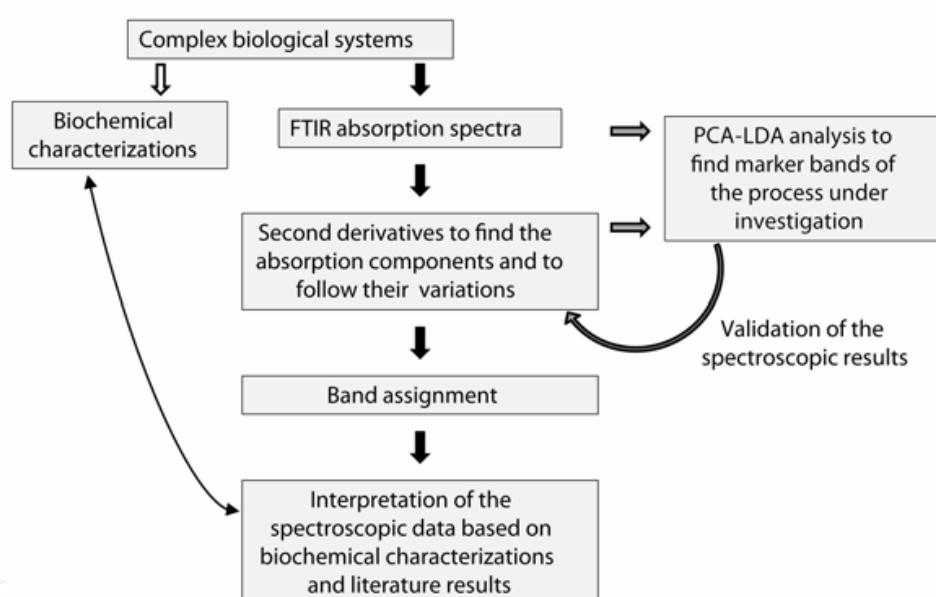


Fig. 1. Scheme of a FTIR approach to study complex biological systems. The measured absorption spectra are analyzed by resolution enhancement approaches, as second derivatives, to resolve the overlapped absorption components and to follow their variations during the process under investigation. To validate the spectroscopic results, a multivariate analysis - such as PCA-LDA - is required. The assignment of the identified marker bands to specific biomolecules involved in the process is the next crucial step. The interpretation of the spectroscopic data should be then confirmed by standard biochemical characterizations.

In this chapter we will first give an overview of FTIR spectroscopy of isolated biomolecules: proteins, nucleic acids, lipids, and carbohydrates. Then, we will illustrate in detail the basis of the multivariate analysis applied to the study of complex biological systems. Finally, we will extend the spectroscopic study on intact cells, focusing our interest on the characterization of embryonic stem cell differentiation.

## 2. FTIR spectroscopy of biomolecules: proteins, nucleic acids, lipids and carbohydrates

The application of FTIR (micro)spectroscopy to the study of biological systems is based on the knowledge of the band assignment of the infrared absorption due to the functional groups of the most important biomolecules. Indeed, proteins, nucleic acids, lipids and carbohydrates have specific absorptions in the mid infrared range, between 4000 and 400  $\text{cm}^{-1}$ . To better illustrate this point, in Figure 2 the IR absorption spectra of model biomolecules are reported and compared with that of intact eukaryotic cells.

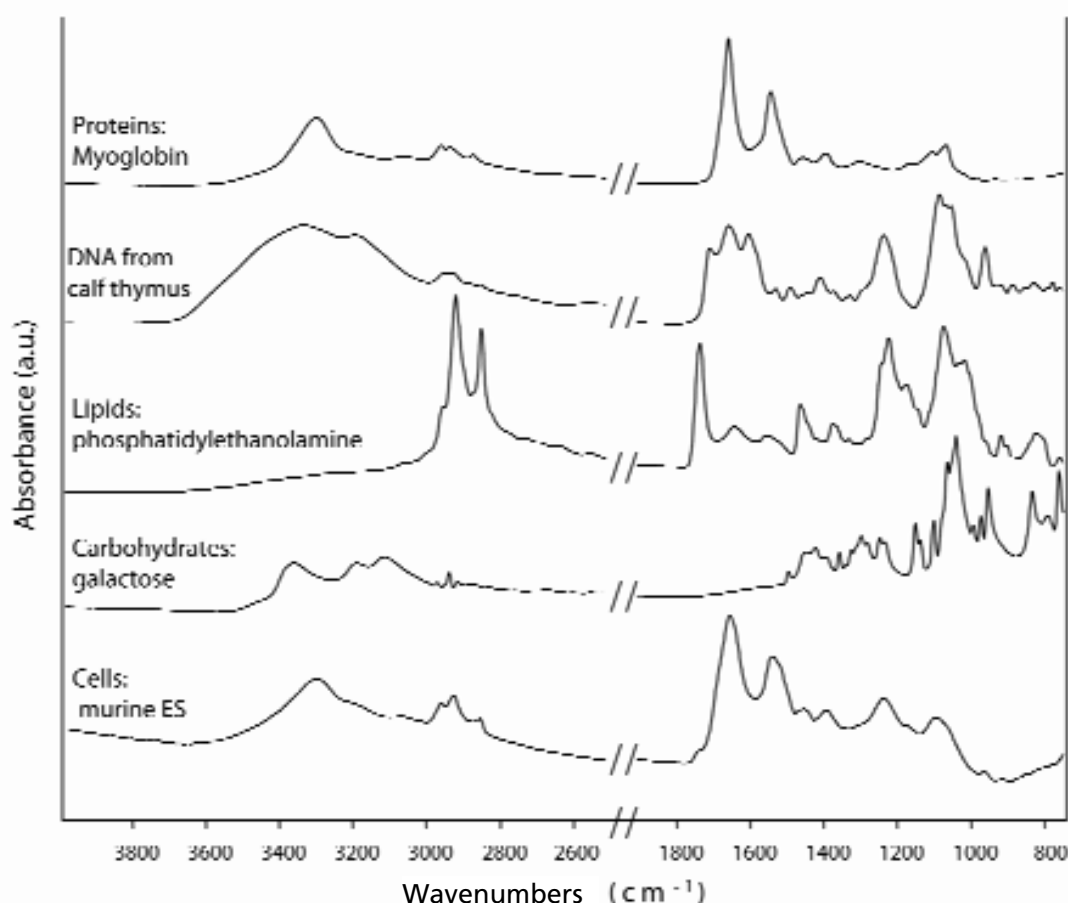


Fig. 2. FTIR absorptions of model biomolecules and intact eukaryotic cells. Myoglobin, calf thymus DNA, phosphatidylethanolamine and galactose are taken as models for protein, nucleic acid, lipid, and carbohydrate IR absorptions respectively. A representative FTIR spectrum of intact murine ES cells is also reported for comparison. Hydrated films of the isolated biomolecules were measured in attenuated total reflection (ATR), while murine embryonic stem (ES) cells were measured in transmission, after dry-fixing for about 30 minutes.

Several spectroscopic studies on complex biological systems are found in literature and became possible since IR spectroscopy is not limited by the physical state of the sample (liquid, solid, etc).

A limiting factor in the infrared characterization of biological molecules in their natural state was initially represented by the strong water absorption in the mid-IR range, where their

internal vibrational modes occur. However, the development of high performing FTIR spectrometers that allow to obtain spectra with an excellent signal to noise ratio and baseline stability enables, nowadays, to subtract the solvent spectrum. Moreover, several strategies – as the use, for instance, of deuterated water – could help to overcome this problem.

For the FTIR study of secondary structures, stability, and aggregation of proteins, particularly useful are the Amide I and the Amide II bands that occur, respectively, in the 1700-1600  $\text{cm}^{-1}$  and 1600-1500  $\text{cm}^{-1}$  spectral regions. The Amide I band, the most used for protein analyses, is mainly due to the C=O stretching vibration of the peptide bond and it is sensitive to the protein secondary structures (Barth, 2007; Barth & Zscherp, 2002; Arrondo & Goni, 1999; Arrondo et al., 1993).

Since FTIR spectroscopy allows to examine also highly scattering samples, proteins can be studied in different environmental conditions, including solutions, hydrated film, and also within intact cells and tissues.

The Amide I of proteins and peptides usually appears as a broad band due to the overlapping of several spectral components arising from the peptide bond absorption in the different secondary structures. On the basis of computational analyses and experimental studies on model compounds - peptides and proteins with known three dimensional structures - it has been possible to assign these components to specific protein secondary structures according to their peak position (Barth, 2007; Barth & Zscherp, 2002; Arrondo & Goni, 1999; Arrondo et al., 1993). For this reason the first critical step on the FTIR analysis of proteins is the identification of the spectral components that contribute to the Amide I envelop, while the second step is the assignment of each component to a specific protein secondary structure. For the first step, two resolution enhancement procedures can be used: the second derivative analysis of the spectra (see below; Susi & Byler, 1986) and the Fourier self-deconvolution (FSD) method (not reported in this chapter, see Arrondo et al., 1993; Kauppinen et al., 1981). Several reviews discuss extensively the Amide I band assignment to protein secondary structures (for instance Barth, 2007; Barth & Zscherp, 2002; Arrondo & Goni, 1999; Arrondo et al., 1993) and here we report only a scheme for protein in not-deuterated solvent: alpha-helices (1660-1648  $\text{cm}^{-1}$ ), beta-sheets (1640-1623  $\text{cm}^{-1}$  and 1695-1674  $\text{cm}^{-1}$ ), turns (1686-1662  $\text{cm}^{-1}$ ), random coils (1657-1642  $\text{cm}^{-1}$ ), and aggregation and protein-protein interactions (1630-1620  $\text{cm}^{-1}$  and 1698-1692  $\text{cm}^{-1}$ ).

Concerning nucleic acids, their IR absorption is very complex and covers a wide range of frequencies (Banyay et al., 2003; Zhizhina & Oleinik, 1972; Tsuboi, 1961). For simplicity, the range of absorption is conventionally divided in different spectral regions. Here, we will briefly illustrate the most studied ones.

The 1800-1500  $\text{cm}^{-1}$  range is mainly due to nucleobase vibrations, sensitive to base stacking and base pairing interactions, while in the 1500-1250  $\text{cm}^{-1}$  range marker bands sensitive to sugar puckering, glycosidic bond rotation and backbone conformation are found. Bands sensitive to nucleic acid backbone conformation occur also between 1250-1000  $\text{cm}^{-1}$ , due to vibrations along the sugar-phosphate chain. This last spectral region is of particular interest for cell biology studies, since a number of marker bands of the different DNA conformations (A, B, and Z) can be found in the spectra, allowing to obtain important insights into the nucleic acid dynamics and functions. For instance, it is known that double stranded DNA exists in two main family forms, the A and the B geometries. In particular, it is known to assume the A form in low relative humidity conditions and in the hybrid with RNA, during transcription (Banyay et al., 2003 and references therein).



Also of particular interest is the spectral range between 1000 and 800  $\text{cm}^{-1}$  where bands due to the different sugar puckering modes (S-N types) are found. Since these are sensitive to changes in the DNA sugar conformation induced by cytosine methylation (Banyay & Graslund, 2002), the analysis of this range could be relevant for biological studies considering the extent of DNA methylation in cells. Theophanides & Tajmir-Riahi (1985) studied the conformational changes of DNA, identifying IR marker bands of A and B DNA forms thanks to the absorption of their different sugar conformations.

Among the first nucleic acid FTIR studies, it should be mentioned the work of Tsuboi (1961), that studied the secondary structure of DNA in solution, native and denatured, enabling to detect the spectral changes due to the breakdown of its secondary structure.

Being the main components of biological membranes, also the infrared absorption of lipids has been widely characterized (Casal & Mantsch, 1984; Arrondo & Goni, 1998). It originates mainly from molecular vibrations of the hydrophylic head-group and of the hydrophobic hydrocarbon tail. The most studied lipid spectral range is that between 3100-2800  $\text{cm}^{-1}$ , where the acyl chain vibrational modes occur with generally strong bands due to CH<sub>2</sub> and CH<sub>3</sub> stretching modes. The frequencies of these bands are conformation-sensitive and respond to temperature-induced changes of the trans/gauche ratio in acyl chains. In this way, it is possible to study lipid phase transitions and changes in lipid composition (Casal & Mantsch, 1984). For instance, through the analysis of the changes in the CH<sub>2</sub> and CH<sub>3</sub> absorption bands, the stress response induced by protein aggregation in bacterial cells has been monitored in situ (Ami et al., 2009). Furthermore, through the CH stretching region analysis of mouse oocyte FTIR spectra and supported by the ester carbonyl band around 1740  $\text{cm}^{-1}$  (see below), Wood and colleagues (Wood et al., 2008) found that lipids - whose composition within the oocytes drastically changes during maturation stages - could be considered potential markers of oocyte developmental competence.

Interestingly, the methyl group vibrations occurring in the 3100-2800  $\text{cm}^{-1}$  spectral range could also give information on DNA/histone methylation and/or histone acetylation, important issues for epigenetic studies (Ami et al., 2010; O'Connell, 2005).

Furthermore, also of particular interest for lipid studies is the so called interfacial region, between 1750-1700  $\text{cm}^{-1}$ , where the stretching vibrations of the C=O group involved in ester bonds occur. The resulting absorption bands are sensitive to changes in their local environment, such as polarity or hydrogen bonding (Casal & Mantsch, 1984; Arrondo & Goni, 1998).

We should add that FTIR spectroscopy has been also widely applied in several fields of carbohydrate research as it allows, for instance, to study mono and oligosaccharide composition and conformation (Kacurakova & Wilson, 2001) and protein glycosylation (Natalello et al., 2005). To this aim, the most used spectral range is in the fingerprint region between 1200 and 750  $\text{cm}^{-1}$ , whose band peak positions and intensities are specific for every polysaccharide. In particular, the IR response has been found to be highly sensitive to the carbohydrate conformation, to hydrogen bonding, to hydration, to the type of substituent, and to the linkage positions (Kacurakova & Wilson, 2001). Due to the complexity of the carbohydrate IR absorption, the band assignment could be not unequivocal, requiring accurate data analysis and validation of the results.

Noteworthy, of particular interest for cell biology applications is the infrared response of glycogen that in tissues and in single cells displays its spectral signature at specific wavenumbers:  $\approx 1028$ ,  $\approx 1081$ , and  $\approx 1153$   $\text{cm}^{-1}$  (Heraud et al., 2010; Ami et al., 2008; Walsh et al., 2007; Wang et al., 2007; Steller et al., 2006).

### 3. Infrared microspectroscopy applied to the study of intact cells: sample preparation

Coupling an infrared microscope, with all reflecting optics, with a FTIR spectrometer offers the opportunity to study selected areas within the sample under investigation.

Two main different types of infrared microscopy can be used: the first, conventional, allows the collection of the IR absorption spectra from a microvolume within the sample, with a spatial resolution not only due to the diffraction limit of Mid IR light (3-10  $\mu\text{m}$ ), but also to the level of the absorbed light. For these reasons, a spectrum of good quality can be collected by an area larger than 20  $\mu\text{m}$  x 20  $\mu\text{m}$  – when using a nitrogen cooled Mercury Cadmium Telluride (MCT) detector (Orsini et al., 2000). Thanks to the variable aperture of the microscope, it is possible to select a small area of few tens of microns within the sample, enabling the study of intact cells (Tanthanuc et al., 2010; Thumanu et al., 2009; Ami et al., 2008; Wood et al., 2008), tissues (Choo et al., 1996) and whole model organisms, as nematodes (Diomedea et al., 2010; Ami et al., 2004).

Instead, more advanced IR microscopes employ a focal plane array (FPA) detection to collect the IR chemical imaging of the sample, with a spatial resolution improved compared to that of the conventional microscope. The image contrast is determined by the response of the different sample regions to the particular IR wavelengths selected by the user (Kazarian & Chan 2006; Levin & Bhargava 2005; Lewis et al., 1995).

We should mention that infrared measurements can be performed mainly in transmission or in attenuated total reflection (ATR) mode. Typically, measurements on biomolecules and on complex biological systems, such as cells and tissues, are carried out in transmission, employing suitable IR transparent supports, for instance of barium fluoride or zinc selenide. However, sometimes, it could be useful to work in ATR when the samples of interest are highly absorbing or when they cannot be easily transferred onto an opportune infrared support (Walsh et al., 2007; Orsini et al., 2000). Indeed, in ATR measurements the sample is placed in contact with the ATR element (diamond, germanium, etc) characterized by a refractive index higher than that of the sample. In this device, an evanescent wave generates and penetrates into the sample for a path length of the size of a micron (Tamm & Tatulian 1997). In addition, the development of synchrotron light sources further improved the application of FTIR microspectroscopy to cell characterization. Indeed, due to its radiation source 100–1000 times brighter than that of a conventional thermal one, it is possible to collect an infrared absorption spectrum at a higher spatial resolution from a sample area of only few microns. In this way, a synchrotron IR source enables to collect high signal – to –noise ratio spectra of subcellular compartments, providing better insights useful for the study of biological processes within single cells (Miller & Dumas, 2010).

We should recall that - even if samples in different physical states can be examined by FTIR spectroscopy - the sample condition can strongly affect the FTIR spectra. This makes it necessary to standardize sample preparation and data acquisition procedures. Indeed, as discussed above for isolated biomolecules, water absorption - very high in the mid-IR - could represent a limit for FTIR analysis, as it makes difficult to perform measurements *in vivo* that require an aqueous environment. A successful strategy to overcome this problem is the dry fixing procedure: a cell suspension is deposited on an IR transparent support and then dried at room temperature for about 30 minutes, in order to remove excess water that would mask the IR response of the different biological components (Ami et al. 2008). Noteworthy, using Raman spectroscopy - a vibrational technique where the water signal is weak and does not affect the

Raman response of hydrated samples - it has been recently demonstrated that the rapid desiccation of cells doesn't affect their spectroscopic response (Konorov et al., 2010). In particular, using Raman microspectroscopy the authors measured different stages of differentiation of human embryonic stem cells live and dry-fixed, and compared the spectroscopic responses obtained in the two conditions. The relative intensities of the bands due to tryptophan in proteins and to nucleic acid backbone and base vibrations - used as differentiation markers - were found to be the same in living and in dry-fixed cells, allowing to monitor the same temporal patterns during differentiation in the two conditions.

These results strongly indicate that dry-fixing is a suitable method for the study of intact cells by FTIR microspectroscopy.

Furthermore, as recently pointed out by Zhao and colleagues (Zhao et al., 2010), changes due to aging of cells in culture could also interfere with the cell IR response. For this reason, the in-situ spectroscopic characterization of cell processes requires an accurate control of the stage of cell growth in culture, in order to obtain reliable and reproducible results.

#### 4. FTIR second derivative analysis

Considering the complexity of the FTIR spectra of biological systems, it is often necessary to better resolve their absorption bands, often broad and overlapped one to the other, using the so called *resolution enhancement procedures*. To this aim, the second derivative analysis is widely applied to the measured spectra, as described and discussed by Susy & Byler (1986). In this way, the overlapping absorption components in the spectrum are identified as negative bands in the second derivative. This analysis requires spectra with high signal-to-noise ratio and free of vapour absorption, as second derivative band intensity is inversely proportional to the square of the original band half-width, leading to an enhancement of the relative contribution of sharp lines, such as due to noise and vapour.

Noteworthy, changes in the relative contributions of the different spectral components can be accurately monitored through the variations in intensity and peak position of the second derivative spectrum.

#### 5. Multivariate statistical analysis

The multivariate statistical analysis (MVA) is an essential tool which allows to tackle the study of complex phenomena which are in general dependent by more than one statistical variable.

In general, the MVA allows the:

- simultaneous treatment of many variables and observations;
- discovery and visualization of complex associations;
- reduction of number of variables;
- construction of descriptive models;
- classification of data into groups.

Among the numerous potential applications of MVA, we are now stressing the aspects related to variable reduction, descriptive models and data classification, which are generally applied to the analysis of spectroscopic data.

In several cases, we are facing the question to find out which are the distinctive traits (if they exist) among samples of experiments done in different conditions, or at different times.



Each experiment can be repeated many times keeping fixed the experimental conditions. We then define a *group* as a collection of two or more replica of the same experiments. We also define the term *instance* or *observation* to refer to a specific experiment within one group.

On every instance, we perform one or more measurements, which we believe are able to capture the fundamental variation among our groups. In such a way, we can express every instance as a vector composed by all our measurements. The measurement can also be a single one, but intrinsically composed of many variables, for example, an IR spectra, where each wavenumber corresponds to a different variable within the same measurement.

In some cases it is also possible that we do not know, *a priori*, the distinction among different groups, but we would like only to determine if it is possible to classify our experiments into distinct groups.

A very broad range of techniques has been developed to address these issues; they span from the statistical analysis to the machine learning field.

For the purpose of this book, we will focus on those methodologies which are particularly successful in the field of spectroscopy, namely the principal component analysis (PCA) and the linear discriminant analysis (LDA).

In the first part, we will explain the basis of PCA with the role of dimensionality reduction combined with LDA as a method for descriptive analysis and classification.

In the second section, we will briefly describe some other fundamental methodologies frequently used in multivariate statistical analysis.

### 5.1 Principal component analysis (PCA)

We describe here the basis of PCA with the specific aim of reducing the number of variables of our problem.

As already mentioned in the introduction, we can express every observation as a vector composed by all our measurements. For example, suppose we have  $n$  observations, each one defined by a vector  $\mathbf{y}_i$  composed of  $m$  variables, where  $i=1,2,\dots,n$  stands for the  $i$ -th observation.

The matrix of the original data  $\mathbf{Y}$  is then composed by  $n$  rows (the observations) and  $m$  columns (the variables).

We do not need, for the PCA, any information about the group membership of the observations, since no grouping of the observations or partitioning of the variables into subsets is assumed.

By using PCA, our intent is to develop a smaller number of artificial variables (called *principal components*) that will account for most of the variance in the observed variables.

We make the assumption that the original variables are redundant, which means that some variables are correlated to each other.

Considering the linear combination of the original data,  $\mathbf{Z} = \mathbf{A}\mathbf{Y}$ , we want to find the matrix  $\mathbf{A}$  such that the new variables  $\mathbf{Z}$  (the principal components) are uncorrelated. The correlation between variables can be measured using the covariance matrix. Given the

sample mean of the  $m$ -dimensional vector  $\mathbf{y}_i$ ,  $\langle \mathbf{y} \rangle = \frac{1}{n} \sum \mathbf{y}_i$ , an unbiased estimator of the sample covariance matrix is

$$\mathbf{S} = \frac{1}{n-1} \sum (\mathbf{y}_i - \langle \mathbf{y} \rangle)(\mathbf{y}_i - \langle \mathbf{y} \rangle)^T \quad (1)$$

For uncorrelated variables, the off-diagonal values of the sample covariance matrix are zero, that is,  $\mathbf{S}$  is diagonal. The covariance of linearly transformed variables  $\mathbf{Z} = \mathbf{A}\mathbf{Y}$  is equal to  $\mathbf{S}_z = \mathbf{A}\mathbf{S}\mathbf{A}^T$ , where  $\mathbf{S}$  is the sample covariance of the original data  $\mathbf{Y}$  (Rencher, 2002).

Thus, we want to find  $\mathbf{A}$  such that the covariance matrix of the transformed data,  $\mathbf{S}_z$ , is diagonal, which corresponds to find the eigenvectors of the covariance matrix and the corresponding eigenvalues.

The eigenvalues, which coincide with the matrix  $\mathbf{S}_z$ , are the sample variance of the principal components  $\mathbf{Z}$  and are ranked according to their magnitude. The first principal component is then the linear combination with maximal variance (largest eigenvalue). The second principal component is the linear combination with the maximal variance along a direction orthogonal to the first component, and so on (Manly, 2004).

The number of eigenvalues is equal to the number of original variables; however, since the eigenvalues are equal to the variance of the principal components and they are sorted in a decreasing order, the first  $k$  eigenvalues explain a large portion of the variance of the data. Hence, to describe our original dataset we can use only the first  $k$  uncorrelated principal components, instead of the complete set of redundant  $m$  variables. In matrix notation this can be written as  $\mathbf{Z}_k = \mathbf{A}_k\mathbf{Y}$  where  $\mathbf{A}_k$  is the eigenvectors matrix truncated to the  $k$ -th eigenvector, and  $\mathbf{Z}_k$  is the matrix of the first  $k$  principal components.

To choose how many principal components should be retained in order to summarize our data, we can use several strategies (Eriksson et al., 2006; Rencher, 2002). For example, one way commonly used is to retain sufficient components to explain a given total percentage of the variance, e.g. 90% (Eriksson et al., 2006; Manly, 2004).

The principal components obtained in this way can be used as a non redundant input for another analysis.

## 5.2 Linear discriminant analysis (LDA)

LDA is mainly a supervised technique, that is, it requires the knowledge of the group membership of the observations. Contrary to PCA, we assume that our data are partitioned into  $k$  groups (Fearn, 2002; Rencher, 2002; Fukunaga, 1990).

LDA can have mainly two objectives. First, it can be a descriptive analysis used to describe and explain the differences among the groups. As we will see later, mathematically LDA finds the optimal hyperplane that separates the groups among each other. Or, in other words, it finds the optimal linear combination of the original variables that maximizes the distance among the groups. The transformed observations are called *discriminant functions*. The use of a linear combination implies that each original variable is weighted by a coefficient, which can be used to study the relative importance of the variable in the separation among the groups. A second possible role of LDA is to classify observations into groups. An observation, whose group membership is not known, is evaluated by a discriminant function (already calibrated) and it is assigned to one of the groups at which most likely it belongs (Eriksson et al., 2006; Manly, 2004; Rencher, 2002).

Firstly, we will explain the “several groups discriminant analysis” applied as a descriptive technique, then we will show how it can be used as a classifier.

### 5.2.1 Several groups descriptive discriminant analysis

The initial dataset is an ensemble of multivariate observations partitioned into  $k$  distinct groups (e.g. different experimental treatments, times or conditions). Each of the  $k$  groups

contains  $n_i$  observations, where  $i$  runs from 1 to  $k$  and refers to the  $i$ -th group. The multivariate observation vectors can be written as  $\mathbf{y}_{ij}$  where  $i=1,\dots,k, j=1,\dots,n_i$  ( $i$  is the  $i$ -th group and  $j$  is the  $j$ -th observation). The vector has size  $m$ , which corresponds to the number of variables.

Our goal in LDA is to search for the linear combination that optimally separates our multivariate observation into  $k$  groups.

This can be visualized in the two group case in Figure 3. The new axis (the discriminant function) allows a better separation of the two clouds of points representing the two dimensional observations of two groups.

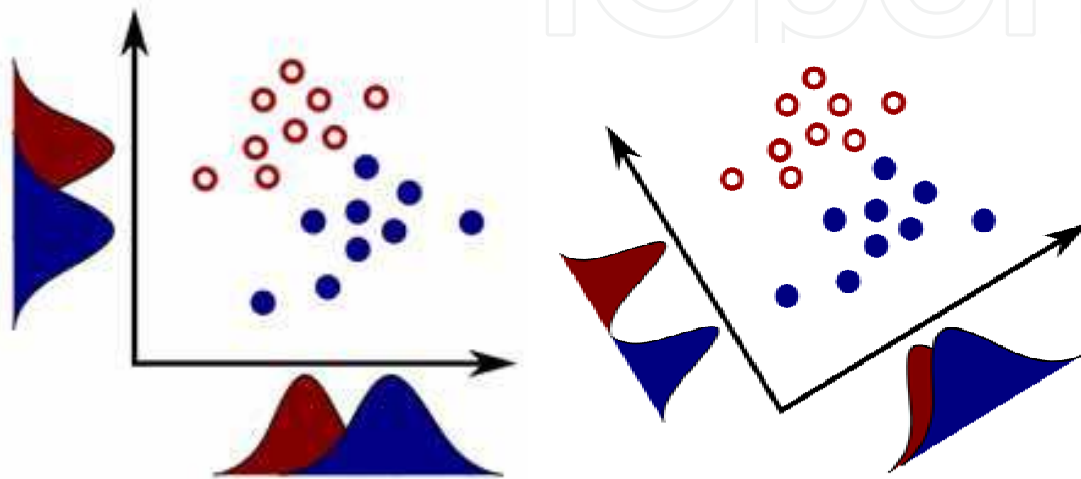


Fig. 3. LDA two group case separation. Rotation of the axis along the direction of the maximal separation between the two groups. After the rotation the two groups can be totally distinguished using only one axis.

The linear transformation of  $\mathbf{y}_{ij}$  is written as  $z_{ij} = \mathbf{w}^T \mathbf{y}_{ij}$

since  $z_{ij}$  is a linear transformation of  $\mathbf{y}_{ij}$ , the mean of the group  $i$  of the transformed data can be written as

$$\langle z_i \rangle = \mathbf{w}^T \langle \mathbf{y}_i \rangle \quad (2)$$

where  $\langle \mathbf{y}_i \rangle$  is the mean of the original variables, obtained as  $\langle \mathbf{y}_i \rangle = \sum \mathbf{y}_{ij} / n_i$

We now introduce the between-groups sum of squares  $\mathbf{B}$  (measure of dispersion among the groups) and the within-group sum of squares  $\mathbf{E}$  (measure of dispersion within one group). First, we define them for the unidimensional case relatively to the untransformed data.

$$\mathbf{B}(\mathbf{y}) = \sum_i n_i (\langle \mathbf{y}_i \rangle - \langle \mathbf{y} \rangle)^2$$

and

$$\mathbf{E}(\mathbf{y}) = \sum_i \sum_j (\mathbf{y}_{ij} - \langle \mathbf{y}_{ij} \rangle)^2$$

where  $\langle \mathbf{y} \rangle = \frac{1}{k} \sum \frac{1}{n_i} \sum \mathbf{y}_{ij}$  is the total average of the data.

Analogously, in the multivariate case (where each observation is constituted by  $m$  variables) we have the two matrices

$$\mathbf{B}(\mathbf{y}) = \sum n_i (\langle \mathbf{y}_i \rangle - \langle \mathbf{y} \rangle) (\langle \mathbf{y}_i \rangle - \langle \mathbf{y} \rangle)^T \quad (3)$$

and

$$\mathbf{E}(\mathbf{y}) = \sum \sum (\langle \mathbf{y}_{ij} \rangle - \langle \mathbf{y}_i \rangle) (\langle \mathbf{y}_{ij} \rangle - \langle \mathbf{y}_i \rangle)^T \quad (4)$$

Finding the optimal linear combination that separates our multivariate observations into  $k$  groups means to find the vector  $\mathbf{w}$  which maximizes the rate between the between-groups sum of squares over the within-groups sum of squares. Using the equation for the mean of the transformed data (eq. 2) into the equations 3 and 4 we can write

$$\lambda = \frac{\mathbf{w}^T \mathbf{B}(\mathbf{y}) \mathbf{w}}{\mathbf{w}^T \mathbf{E}(\mathbf{y}) \mathbf{w}} = \frac{\mathbf{B}(\mathbf{z})}{\mathbf{E}(\mathbf{z})} \quad (5)$$

We want to find  $\mathbf{w}$  such that  $\lambda$  is maximized.

Equation 5 can be rewritten in the form  $\mathbf{w}^T (\mathbf{B}\mathbf{w} - \lambda \mathbf{E}\mathbf{w}) = 0$ ; then we search for all the non trivial ( $\mathbf{w}^T = 0$  is excluded) solutions of this equation and we choose the one which gives the maximum value of  $\lambda$ . This means to solve the eigenvalue problem  $\mathbf{B}\mathbf{w} - \lambda \mathbf{E}\mathbf{w} = 0$  which can be written in the usual form

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{w} = 0 \quad (6)$$

where  $\mathbf{A} = \mathbf{E}^{-1}\mathbf{B}$

The solutions of equation 6 are the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_m$  associated to the eigenvectors  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ . The solutions are ranked for the eigenvalues  $\lambda_1 > \lambda_2 > \dots > \lambda_m$ . Hence, the first eigenvalue  $\lambda_1$  corresponds to the maximum value of eq. 5.

The discriminant functions are then obtained considering only the first  $s$  positive eigenvalues and multiplying the original data by the eigenvectors

$$\mathbf{z}_1 = \mathbf{w}_1^T \mathbf{Y}, \quad \mathbf{z}_2 = \mathbf{w}_2^T \mathbf{Y}, \dots, \mathbf{z}_s = \mathbf{w}_s^T \mathbf{Y}$$

Discriminant functions are uncorrelated but not orthogonal since the matrix  $\mathbf{A} = \mathbf{E}^{-1}\mathbf{B}$  is not symmetric.

In many cases the first two or three discriminant functions account for most of  $\lambda_1 + \lambda_2 + \dots + \lambda_s$ . This allows to represent the multivariate observations as 2 or 3 dimensional points which can be plotted on a scatter plot. These plots are particularly helpful to visualize the separation of our observations into the different groups. Moreover, we can deduce, looking at the scatter plot, the meaning of a given discriminant function, i.e. we can associate the discriminant function to a given property of the analyzed system.

The weighting vectors  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_s$  are called *unstandardized discriminant function coefficients* and give the weight associated to each variable on every discriminant function.

If the variables are on very different scales and with different variance, to assess the importance of each variable in the group separation the standardized discriminant functions

can be used. The standardization is done by multiplying the unstandardized coefficients by the square root of the diagonal element of the within-group covariance matrix. Another way to assess the variable importance is to look at the correlation between each variable and the discriminant function. These correlations are called *structure* or *loading* coefficients. However, it has been shown (Rencher, 2002) that these parameters are intrinsically univariate, and they only show how a single variable contributes to the separation among groups, without taking into account the presence of the other variables.

5.2.2 LDA as classification method

The discriminant analysis can be applied to a given ensemble of data to produce a set of discriminant functions as described in the previous section. Afterwards, this model can be used to classify new observations into the most probable groups. From this point of view the linear discriminant analysis becomes a predictive tool, since it is able to classify observations whose group membership is unknown (Eriksson et al., 2006; Rencher, 2002). In the same way we can test the discrimination ability of our LDA model by a procedure called “*re-substitution*” (Rencher, 2002). This method consists of producing an LDA model using our dataset (i.e. finding the optimal  $w$ ). Then, each observation vector is re-submitted to the classification function ( $z_{ij} = w^T y_{ij}$ ) and assigned to a group. Since we know the group membership of the submitted vector, we can count the number of observations correctly classified and the number of observations misclassified.

Then, we can estimate the apparent classification rate as the number of correctly classified observations over the total number of observations. This is summarized in a classification table or confusion matrix. As an example, given  $N$  observations,  $n_1$  belongs to the group 1 and  $n_2$  belongs to the group 2.  $C_{11}$  is the total number of observations correctly classified in group 1 and  $C_{12}$  is the total number of data misclassified in group 2. Similarly,  $C_{22}$  is the total number of observations correctly classified in group 2 and  $C_{21}$  is the number of misclassified in group 1.

The confusion matrix becomes then:

Actual group	Predicted group	
	1	2
1	$C_{11}$	$C_{12}$
2	$C_{21}$	$C_{22}$

And the accuracy (the apparent classification rate (acr)) is computed as

$$acr = \frac{C_{11} + C_{22}}{n_1 + n_2}$$

In general, in evaluating the accuracy of a model, we have then to distinguish between two types of accuracy: the fitting accuracy and the prediction accuracy (Eriksson et al., 2006; Bishop, 1995).

The fitting accuracy is the ability to reproduce the data, namely, how the model is able to reproduce the data that were used to build the model. This corresponds to the apparent classification rate, and it is obtained using the *re-substitution* procedure. The data used to build the model are called *training set*.



The prediction accuracy is the ability to predict the value or the class of an observation, which was not included in the construction of the model. This kind of accuracy is often referred as the ability of the model to generalize. The data used to measure this accuracy are called *test set*. The prediction accuracy can be called actual classification rate. This is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice.

To have an estimation of the actual classification rate, two main procedures can be applied: the hold-out and *cross-validation* (Eriksson et al., 2006).

In the hold-out, the dataset is divided into two partitions, one partition is used to develop the model (e.g the discriminant functions) and the second partition is given as input to the model. The first partition is usually called training set or calibration set, while the second partition is the validation set (Bishop, 1995).

When the number of observations is small, the *cross-validation* is usually preferred over the *hold-out*. The basic idea of the cross-validation procedure is to divide the entire dataset into  $L$  disjoint sets.  $L-1$  sets are used to develop the model (i.e. this is the calibration set on which the discriminant functions are computed) and the omitted portion is used to test the model (i.e. the validation set given as input to the model). This is repeated on for all the  $L$  sets and an average result is obtained.

### 5.3 Principal component - linear discriminant analysis (PCA-LDA)

A powerful analysis tool is the combination of the principal component analysis with the linear discriminant analysis (Fearn, 2002). This is particularly helpful when the number of variables is large. In particular, if the number of observations ( $N$ ) is less than the number of variables ( $m$ ) - specifically  $N-1 < m$  - the covariance matrix is singular and can not be inverted. We then need to find a way to reduce the number of variables, for example by using the PCA (Rencher, 2006; Jonathan et al., 1996).

This procedure has been widely used for several problems in different fields (Ami et al., 2008; Rezzi et al., 2007; Skrobot et al., 2007; Walsh et al., 2007; Pereira et al., 2006; Héberger et al., 2003; Fearn, 2002).

In particular, a low rate of  $(N-1)/m$  happened normally in spectroscopy, where the number of observations ( $N$ ) is usually  $< 10^2$  and the number of variables ( $m$ ) is typically within  $10^2$  to  $10^3$ . Let's take into account the same situation described for many group linear discriminant analysis. The original dataset is an ensemble of multivariate observations which is partitioned into  $k$  distinct groups. Again, we want to find the discriminant functions which optimally separate our multivariate observation into the  $k$  groups. Then, the discriminant functions can be used to identify the most important variables in terms of ability of distinguishing among the groups.

Thus, first the original dataset is submitted to PCA to reduce the number of variables, subsequently the reduced dataset is analyzed using LDA.

### 5.4 Other multivariate techniques

In the following section, we will briefly illustrate other multivariate statistical approaches, relevant for the spectroscopic studies reported in this chapter.

#### 5.4.1 Multivariate Linear Regression (MLR)

MLR can be used to model a linear relationship among a numerical variable  $z$  and one or more independent variables  $Y$  (Manly, 2004).  $Y$  is the usual matrix already introduced,

composed by  $n$  rows corresponding to observations and  $m$  columns corresponding to independent variables. The MLR is based, as many other statistical techniques, on the generalized linear model  $\mathbf{z} = \mathbf{Y}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , where  $\boldsymbol{\beta}$  is a matrix containing the parameters to be estimated,  $\boldsymbol{\varepsilon}$  is a matrix which models the errors or noise.

The coefficients beta are usually estimated using the ordinary least square, which consists of minimizing the sum of the squared differences of the  $n$  observed  $y$ 's from their modeled values. Mathematically, the optimal values of beta are obtained by  $\boldsymbol{\beta} = (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{z}$ . To apply the least square method we must have  $n - 1 > m$ , otherwise the matrix  $\mathbf{Y}^T \mathbf{Y}$  is singular and can not be inverted. Moreover, none of the independent variables must be a linear combination of any other (muticollinearity) (Eriksson et al., 2006; Manly, 2004).

#### 5.4.2 Partial Least Square (PLS)

The goal of PLS regression is to predict  $\mathbf{Z}$  from  $\mathbf{Y}$  and to describe their common structure. When the number of variables is large compared to the number of observations,  $\mathbf{Y}$  is likely to be singular and the regression approach is no longer feasible (i.e., because of multicollinearity) (Eriksson et al., 2006). Several approaches have been developed to cope with this problem. One approach is to eliminate some predictors (e.g., using stepwise methods); another one, called *principal component regression*, is to perform a PCA of the  $\mathbf{Y}$  matrix and then use the principal components (i.e., eigenvectors) of  $\mathbf{Y}$  as regressors on  $\mathbf{Z}$ .

The problem is then of choosing an optimum subset of predictors that gives the best regression. One possibility is to choose the first  $k$  principal components; however, these components are obtained to best explain  $\mathbf{Y}$  rather than  $\mathbf{Z}$ , and so, nothing guarantees that they are also relevant for  $\mathbf{Z}$ . In PLS we seek the components from  $\mathbf{Y}$  that are relevant also for  $\mathbf{Z}$ . In particular, PLS regression performs a simultaneous decomposition of  $\mathbf{Y}$  and  $\mathbf{Z}$  into principal components with the constraint that the components explain as much as possible the covariance between  $\mathbf{Y}$  and  $\mathbf{Z}$  (Rencher, 2002).

#### 5.4.3 Factor Analysis (FA)

Factor analysis is a statistical method used to discover if the observed variables can be explained in terms of a much smaller number of variables called factors. It is closely related to PCA in that they both try to reduce the redundancy among the variables by using a smaller number of factors (or principal components in PCA); however it has some important differences: i) in PCA the components are defined as linear combinations of the original variables while in FA the original variables are linear functions of the factors; ii) in PCA we seek to explain the total variance, while in FA we attempt to reproduce the covariance; iii) in PCA essentially no assumptions are required while in FA some fundamental assumptions are defined; iv) the principal components are unique, whereas the factors can be rotated. By rotating your factors you attempt to find a factor solution that is equal to that obtained in the initial extraction but which has the simplest interpretation. This last point is one of the main advantage of the FA over PCA, if our goal is to find and describe the underlying factors of the data. On the other hand, if we are simply searching for a smaller number of variables as input for another analysis, the PCA is preferred (Manly, 2004; Rencher, 2002; Bryant & Yarnold, 1994).

#### 5.4.4 Cluster Analysis (CA)

Cluster analysis is a procedure used to partition the data into groups so that the most similar observations are assigned in the same cluster and clusters are dissimilar to each other (Manly, 2004).

CA is an unsupervised technique, that is, the group membership of the observations (and often the number of groups) is not known in advance.

Since we are trying to group similar observations, a measure of similarity or dissimilarity is required. The most common distance functions are: i) the Euclidean distance; ii) the Manhattan distance; iii) the Mahalanobis distance; iv) the maximum norm.

Several types of clustering algorithms have been developed. Based on the procedure they use, they can be divided into three main groups: hierarchical, partitional and density-based clustering.

Hierarchical clustering algorithms are sequential. They can be agglomerative or divisive. The agglomerative clustering starts with all observations placed in different clusters and in each step an observation or a cluster of observations are merged into another cluster. The divisive method starts with one single cluster containing all observations and then it divides the cluster into two sub-clusters at each step.

The partitional algorithm assigns the observations to a set of clusters without using hierarchical approaches. One of the most used non-hierarchical approach is the k-means clustering.

The density-based clustering seeks to search for region of high density without any assumption about the shape of the cluster.

#### 5.4.5 Artificial Neural networks (ANN)

The artificial neural networks are mathematical models that were developed in analogy to a network of biological neurons (Krogh, 2008). In the brain, the highly interconnected network of neurons communicates sending electric pulses through the neural wiring of axons, synapses and dendrites. Mathematically, a neuron can be modeled as a switch that receives a series of values as input and produces an output consisting of a weighted sum of the input vectors eventually filtered by a function  $f$ . Many neurons can be combined to create more complex networks. Depending on the type of neurons and how the neurons are connected to each other, different kinds of neural networks can be created. The most common type of neural network is the feed-forward neural network, in which neurons are grouped into layers, each neuron of a layer is connected to all the neurons of the next layer and the information flows from the input to the output without loops. For a comprehensive description of neural networks and their applications see Haykin (1999) and Bishop (1995).

### 6. Applications of vibrational spectroscopies to the study of stem cell differentiation

Stem cells (SCs) are self-renewing cells characterized by the capacity to differentiate into a wide range of specialized cells. Two main types of SCs exist: embryonic and adult.

Embryonic stem cells (ESCs) are derived from cells of an embryo - the inner cell mass of blastocyst - and are considered the most versatile type of SCs because they have the unique ability to retain the developmental capacity of generating all functional adult cell types (Thomson et al., 1998; Evans & Kaufman, 1981).

Adult SCs are found among differentiated cells of a tissue or organ that can renew itself and can differentiate into specialized cell types of the same tissue or organ where they reside (the *stem cell niche*). Indeed, their major role is to maintain and repair the tissue in which they are found. Adult SCs have been identified in many organs and tissues, including bone marrow, brain, liver, skeletal muscle, and skin (Verfaillie, 2002; Peterson & Davidson, 2000). As discussed in the introduction, Raman and FTIR microspectroscopies are successfully applied in stem cell research (Chan & Lieu, 2009; Heraud & Tobin, 2009). In the following pages, we will first report some studies that illustrate the potential of these vibrational tools to characterize embryonic stem cells. We will then illustrate a few research works on adult stem cells, that we believe could be relevant for a better understanding of stem cell biology.

### 6.1 Embryonic stem cells

FTIR and Raman microspectroscopies allow to detect rapidly and in a non invasive way biochemical changes during ES cell differentiation, providing unique markers for the identification in-situ of SCs differentiation status.

One of the earliest studies aimed at the spectroscopic characterization of embryonic stem (ES) cells is that of Notingher and colleagues, who applied Raman microspectroscopy to monitor the murine ES cell differentiation process, spontaneous and via embryoid body (EB) formation (Notingher et al., 2004 a and b). The authors showed that undifferentiated, spontaneously differentiated, and EB differentiated murine ES cells exhibit unique Raman markers that, in association with PCA, could be used to identify the differentiation state of the ES cells. In particular, it was found that the most significant differences could be attributed to cell RNA content that in undifferentiated cells was higher than in differentiated ones, a result that suggests to the authors that differentiating ES cells use the pool of dormant mRNA to produce new specific proteins of the new phenotype. Indeed, as the ES cells start to differentiate toward various phenotypes, the translation of mRNA increases, as indicated by the decrease in the ratio between the areas of the  $813\text{ cm}^{-1}$  RNA peak and the phenylalanine peak at  $1005\text{ cm}^{-1}$ , to reach values similar to those found in fully differentiated cells, after 16-20 days of differentiation. These results indicated that RNA and protein peaks in the Raman spectra of murine ES cells can be used as a differentiation marker, with important applications for the development of engineered tissues.

Raman microspectroscopy, coupled with the multivariate PCA-LDA analysis, was also applied to explore the possibility to discriminate between undifferentiated human ES cells and their cardiac derivatives (Chan et al, 2009). Indeed, unlike other cell lineages, cardiomyocytes lack specific surface markers required for their physical identification and separation, making desirable the development of new analytical tools. In this work, the authors were able to detect spectroscopic signatures of ES cells and of their cardiac derivatives, mainly involving RNA and protein content. In particular, the authors found that undifferentiated cells were characterized by a more elevated mRNA level than differentiated cells, resulting from their different active cell cycles, as suggested by the different intensity of the peak at  $811\text{ cm}^{-1}$  (phosphodiester bond) observed in the two cases. Interestingly, their results were in agreement with those obtained by Notingher and colleagues on murine stem cells (Notingher et al., 2004 a and b) - as discussed above - and by Schulze and colleagues (2010) in a Raman study of the spontaneous differentiation process of human ES cells. Noteworthy, they also investigated the effect of laser exposure on cells, in order to verify the non-invasiveness of the spectroscopic method. Indeed, they demonstrated that the laser irradiation does not



compromise cell pluripotency, as it didn't affect the expression of the human ES cells transcription factor OCT4, required to sustain ES cell self renewal. Moreover, no effects on cell morphology and cell proliferation were detected.

Of great interest is the study of Heraud and colleagues (2010) that employed FTIR microspectroscopy with focal plane array detection to characterize human ES cell differentiation directed toward specific cell lineages, namely mesendoderm and ectoderm. Well defined spectral differences – confirmed also by partial least squares discriminant analysis (PLS-DA) and artificial neural network analysis (ANN) – were detected among the three different cell populations, mainly involving the lipid and the glycogen bands (respectively at  $2920\text{ cm}^{-1}$  and at  $1155\text{ cm}^{-1}$ ), whose intensities were found to be higher in the undifferentiated than in the differentiated cell populations. The results demonstrated that FTIR signatures can be used to successfully discriminate between human stem cells and their differentiated progenies, even at early stages of differentiation.

Another application of FTIR microspectroscopy, coupled with PCA and unsupervised hierarchical cluster analysis (UHCA), was aimed at identifying specific marker bands of murine ES cell differentiation toward neural cell types (Tanthanuc et al., 2010). In particular, by applying focal plane array detection and synchrotron based FTIR microspectroscopy, the authors were able to find significant differences between undifferentiated and differentiated cells, mainly in spectral regions due to lipid and protein absorptions. In particular, they observed a dramatic increase of the acyl chain  $\text{CH}_2$  symmetric and asymmetric stretching modes – around  $2850\text{ cm}^{-1}$  and  $2920\text{ cm}^{-1}$  respectively – during the differentiation process, increment possibly related to changes in membrane lipids responsible for neural cell differentiation and signal transduction. This result has been also confirmed monitoring the lipid carbonyl band around  $1740\text{ cm}^{-1}$ , whose peak position and intensity were observed to change during differentiation. Furthermore, important changes in protein secondary structures were detected and in particular the differentiated cells appeared to be characterized by a higher content of alpha-helix proteins than undifferentiated cells. The authors explained this result as due to the increased expression of alpha-helix rich proteins of the cytoskeleton, as tubulin and actin, important for the establishment of neural structure and function.

We applied FTIR microspectroscopy – supported by PCA-LDA analysis – to characterize *in situ* the early stages of murine ES cell spontaneous differentiation (Ami et al., 2010; Ami et al., 2008). We found that significant changes in nucleic acid, protein and lipid content occurred during the differentiation process.

In Figure 4 we reported the second derivative spectra of ES cells at different maturation stages (from undifferentiated to 14 days of differentiation).

As illustrated in the Figure, we first found that undifferentiated ES cells were characterized by a RNA content higher than differentiating cells, in agreement with what reported by Notingher and colleagues with Raman microspectroscopy (Notingher et al., 2004 a and b), as previously discussed. Moreover, we monitored the formation of the DNA/RNA hybrid through the simultaneous presence of the three components respectively around  $954\text{ cm}^{-1}$  (CC stretching of DNA backbone), at  $914\text{ cm}^{-1}$  (ribose ring) and at  $899\text{ cm}^{-1}$  (deoxyribose), after 4-7 days of differentiation. These results, indicating that the transcription activity for the new phenotype was taking place in that temporal range, were further supported by changes in the secondary structures of the whole protein content, likely due to the emergence of the new phenotype. Indeed, again starting from 4-7 days of differentiation, we observed an increase of alpha-helix and beta-turn components, respectively at  $1658\text{ cm}^{-1}$



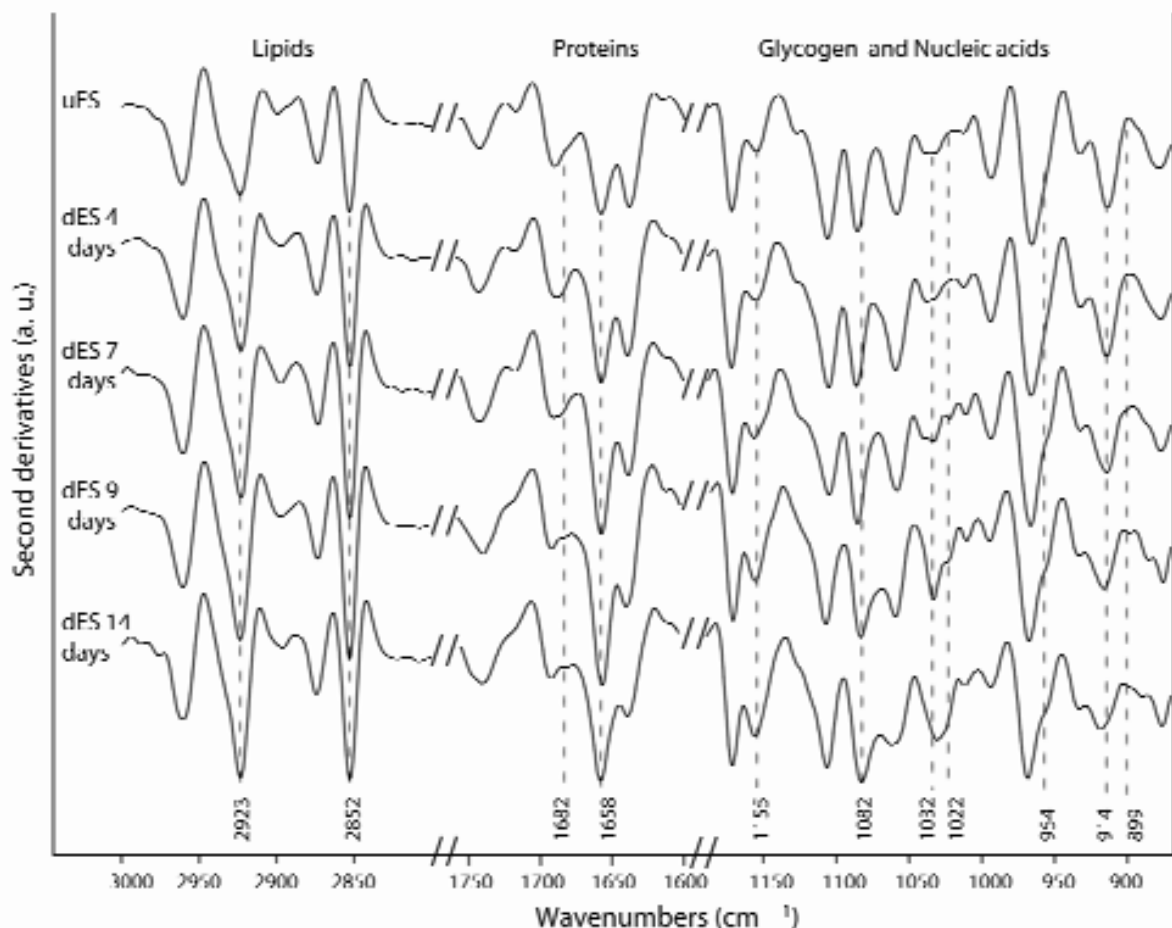


Fig. 4. Murine embryonic stem cell differentiation monitored by FTIR microspectroscopy. The second derivative of FTIR spectra of murine ES cells, undifferentiated (uES) and spontaneously differentiated (dES), are reported in three different spectral regions: i) 3000-2800  $\text{cm}^{-1}$  mainly due to lipid acyl chains; ii) 1750-1600  $\text{cm}^{-1}$  where protein amide I band occurs; iii) 1200-800  $\text{cm}^{-1}$ , mainly due to glycogen and nucleic acid absorptions (see text). Spectra have been normalized at the tyrosine band around 1515  $\text{cm}^{-1}$  and reported after magnification in each region, for the presentation of the data.

and 1682  $\text{cm}^{-1}$ , that suggested that the expression of proteins typical of cardiomyocyte precursors was taking place. Indeed, it is known that these cells are rich in alpha-myosin, a protein belonging to alpha-helix fold and that they are characterized by the formation of gap junctions (Oyamada et al., 1996), whose main protein components are connexins, containing again alpha-helix structures and an important percentage of beta-turns. To support our hypothesis - confirmed by cytochemical analysis - after the "switch" of the new phenotype we also observed the emergence of IR bands due to glycogen at 1155  $\text{cm}^{-1}$ , 1081  $\text{cm}^{-1}$  and between 1035 and 1020  $\text{cm}^{-1}$ , typical of cardiomyocytes (Pasumarthi & Field 2002).

Also dramatic changes in lipid absorption were detected during ES cell differentiation. In particular, an increase of the CH<sub>2</sub> vibrational modes at 2923  $\text{cm}^{-1}$  and at 2852  $\text{cm}^{-1}$  was monitored, starting as soon as the differentiation process was taking place, and up to the end of our investigation (9-14 days). These results indicated that significant changes in lipid composition occurred, suggesting that the new phenotype was characterized by new membrane properties.

The spectroscopic results were then validated by PCA-LDA analysis that allowed to obtain an excellent segregation of the data into five separated clusters, each corresponding to a specific differentiation stage, as reported in Figure 5. Moreover, this analysis enabled us to identify in the spectrum the wavenumbers that contributed to the largest inter-spectral variance during the differentiation process, and - in agreement with the direct inspection of the spectral data - they were found to be due to protein and nucleic acid components.

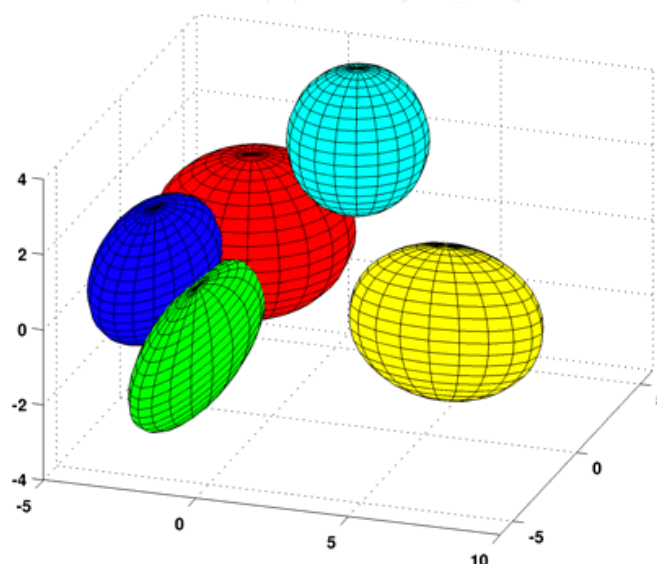


Fig. 5. PCA-LDA analysis of murine ES cell differentiation. The clustering of FTIR absorption spectra - from 1800 to 800  $\text{cm}^{-1}$  - as 3D score plot is shown. Data for undifferentiated cells (red) and at 4 (blue), 7 (green), 9 (light blue), and 14 (yellow) days of differentiation have been analysed. The ellipsoid semi-axes correspond to two standard deviations of the data.

## 6.2 Adult stem cells

An interesting investigation performed by synchrotron based FTIR microspectroscopy - coupled with principal component analysis (PCA) - enabled to discriminate in bovine cornea among SCs, transit-amplifying (TA) and terminally differentiated (TD) cells (German et al., 2006). Measuring the absorption spectra of individual cells in cryosections, the authors found significant spectral differences among the three different cell types, with only a slight overlap between SC and TA cells. The most important differences mainly involved changes in spectral components due to nucleic acid absorptions, like the RNA band at 1120  $\text{cm}^{-1}$  and the phosphate band around 1080  $\text{cm}^{-1}$ . Moreover, the authors found that TD cells formed a well separated and homogeneous population with spectral features closer to TA cells than to SCs. As expected, the spectral response of the terminal differentiation state is characterized by important changes in nucleic acid and protein content, being associated with a loss of proliferative ability and the production of proteins associated with the new phenotype.

A FTIR characterization of human corneal epithelium performed by the same research group (Bentley et al., 2007) confirmed the previous results obtained on bovine cornea. Also in this case, the authors were able to discriminate among SC, TA and TD cells, finding again important changes mainly in nucleic acid and protein content. In particular, the main spectral differences between TA and TD cells were found to involve the protein secondary

structures and the RNA expression, as discussed above. Noteworthy, in the two works, the authors detected small subpopulations of cells within the corneal epithelium SC niche with TA cell like characteristics, strongly suggesting that the TA cells are newly generated prior to their migration.

Furthermore, the entire tissue architecture was investigated using IR spectral imaging that enabled to localize and better characterize SCs (Nakamura et al., 2010). By this approach further details on the differences among SC, TA and TD cells were obtained, confirming that nucleic acid response - between 1425 and 900  $\text{cm}^{-1}$  - accounts for the most significant differences among the three types of cell populations. Important changes in protein content, between 1800 and 1480  $\text{cm}^{-1}$ , have been also detected in the examined cell types, as expected considering their different functions. Interestingly, the most discriminating spectral features of SCs were associated to DNA and RNA conformations, as indicated by the bands at 1225  $\text{cm}^{-1}$  and at 1080  $\text{cm}^{-1}$  respectively, whereas IR bands due to proteins and lipids, respectively at 1558  $\text{cm}^{-1}$  and at 1728  $\text{cm}^{-1}$ , allowed to discriminate between TA and TD cells.

Of particular relevance is also the work of Walsh and colleagues, where synchrotron FTIR microspectroscopy - supported by PCA-LDA analysis - was applied to characterize the different cell types, derived from stem cells, along the length of gastrointestinal tract, one of the most regenerative human tissue (Walsh et al., 2009). Through IR image maps with the related IR spectra collected from tissue sections at the single cell level, the authors detected spectral changes in the differentiation states along the gastrointestinal tract - with common features in related cell types - mainly involving DNA conformational changes, with one of the most important spectral marker at 1080  $\text{cm}^{-1}$ , due to phosphate vibrational mode. These results were further confirmed by PCA-LDA multivariate analysis, whose crucial role has been critically highlighted in Walsh' work. Indeed, this analysis allowed to identify as the most contributory wavenumbers those due to the phosphate mode absorptions, partly associated to protein phosphorylation. Overall, these results suggested to the authors that DNA conformational changes could be considered a significant stemness markers in gastrointestinal crypts.

## 7. Conclusive remarks

The examples reported in this chapter highlight the great potential of spectroscopic approaches providing new insights in stem cell biology. In particular, FTIR microspectroscopy is a powerful tool that enables to obtain - in a non invasive way - a chemical fingerprint of the cell types, giving information on the overall changes in the macromolecular content occurring during a biological event. In this way, this approach allows to assess in-situ the differentiation status of the cells through the identification of specific marker bands. Moreover, the time evolution of these bands enabled to follow the progress of the process by the simultaneous monitoring of the most important cellular components, as nucleic acids and proteins.

We should underline that the successful application of the spectroscopic approach requires the use of an appropriate multivariate analysis to validate the spectral data and to identify the marker bands of the process under investigation. The integration with the established biochemical methods is, of course, an important requisite to understand the biological significance of the spectroscopic results.

As a final comment, we would also like to point out that FTIR and Raman spectroscopic approaches, indeed, might offer preliminary tools - rapid and inexpensive - to obtain useful

information on complex systems, in order to design conclusive biological experiments. As discussed above, these techniques allow to characterize the temporal correlation of biological events that occur simultaneously in a complex system, a result not easily tackled by the standard biochemical methods.

## 8. Acknowledgements

D. A. is indebted to Fondazione IRCCS Policlinico San Matteo, Pavia (I) for the supporting scholarship. S.M. D. acknowledges the financial support of the FAR (Fondo di Ateneo per la Ricerca) of the University of Milano-Bicocca (I). The authors wish to thank Prof. Carlo Alberto Redi and its research group at the University of Pavia (I) for the research collaboration on embryonic stem cell differentiation. We are grateful to Dr Carla Smeraldi for the language revision of this chapter.

## 9. References

- Ami D., Natalello, A., Zullini, A., Doglia, S.M. (2004). Fourier transform infrared microspectroscopy as a new tool for nematode studies. *FEBS Letters*, 576, 297-300, ISSN: 0014-5793.
- Ami, D., Natalello, A., Gatti-Lafranconi, P., Lotti, M., Doglia, S.M. (2005). Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Letters*, 579, 3433-3436, ISSN: 0014-5793.
- Ami, D., Neri, T., Natalello, A., Mereghetti, P., Doglia, S.M., Zanoni, M., Zuccotti, M., Garagna, S., Redi, C.A. (2008). Embryonic stem cell differentiation studied by FT-IR spectroscopy. *Biochimica et Biophysica Acta*, 1783, 98-106, ISSN: 0167-4889.
- Ami, D., Natalello, A., Schultz, T., Gatti-Lafranconi, P., Lotti, M., Doglia, S.M., de Marco, A. (2009). Effects of recombinant protein misfolding and aggregation on bacterial membranes. *Biochimica et Biophysica Acta*, 1794, 263-269, ISSN: 1570-9639.
- Ami, D., Natalello, A., Mereghetti, P., Neri, T., Zanoni, M., Monti, M., Doglia, S.M., Redi, C.A. (2010). FT-IR spectroscopy supported by PCA-LDA analysis for the study of embryonic stem cell differentiation. *Spectroscopy-Biomedical Applications*, 24, 89-97, ISSN: 0712-4813.
- Arrondo, J.L.R., Muga, A., Castresana, J., Goni, F.M. (1993). Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy. *Progress in Biophysics and Molecular Biology*, 59, 23-56, ISSN: 0079-6107.
- Arrondo, J.L.R. & Goni, F.M. (1998). Infrared studies of protein-induced perturbation of lipids in lipoproteins and membranes. *Chemistry and physics of lipids*, 96, 53-68, ISSN: 0009-3084.
- Arrondo, J.L.R., & Goni, F.M. (1999). Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Progress in Biophysics and Molecular Biology*, 72, 367-405, ISSN: 0079-6107.
- Baker, M.J., Gazi, E., Brown, M.D., Shanks, J.H., Gardner, P., Clarke, NW. (2008). FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer. *British Journal of Cancer*, 99, 1859 - 1866, ISSN: 0007-0920.
- Banyay, M. & Graslund, A. (2002). Structural effects of cytosine methylation on DNA sugar pucker studied by FTIR. *Journal of Molecular Biology*, 324, 667-676, ISSN: 0022-2836.



- Banyay, M., Sarkar, M., Graslund, A. (2003). A library of IR bands of nucleic acids in solution. *Biophysical Chemistry*, 104, 477–488, ISSN: 0301-4622.
- Barth, A. & Zscherp, C. (2002). What vibrations tell us about proteins. *Quarterly Reviews of Biophysics*, 35, 369–430, ISSN: 0033-5835.
- Barth, A. (2007). Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta*, 1767, 1073–1101, ISSN: 0005-2728.
- Bentley, A.J., Nakamura, T., Hammiche, A., Pollock, H.M., Martin, F.L., Kinoshita, S., Fullwood, N.J. (2007). Characterization of human corneal stem cells by synchrotron infrared micro-spectroscopy. *Molecular Vision*, 13, 237-242, ISSN: 1090-0535.
- Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, ISBN:0198538642, New York, NY, USA.
- Bryant, F.B. & Yarnold, P.R. (1994). Principal components analysis and exploratory and confirmatory factor analysis, In: *Reading and understanding multivariate analysis*, Grimm & Yarnold (Eds.), American Psychological Association Books, ISBN: 978-1-55798-273-5, Washington DC, USA.
- Casal, H.L. & Mantsch, H.H. (1984). Polymorphic phase behaviour of phospholipid membranes studied by infrared spectroscopy. *Biochimica et Biophysica Acta*, 779, 381-401, ISSN: 03044157.
- Chan, J.W., Lieu, D.K., Huser T., Li R.A. (2009). Label-free, non-invasive spectroscopic separation of human embryonic stem cells (hESCs) and their cardiac derivatives. *Analytical Chemistry*, 81, 1324-1331, ISSN: 0003-2700.
- Chan, J.W. & Lieu, D.K. (2009). Label-free biochemical characterization of stem cells using vibrational spectroscopy. *Journal of Biophotonics*, 2, 656–668, ISSN: 1864-063X.
- Choo, L.P., Wetzel, D.L., Halliday, W.C., Jackson, M., LeVine, S.M, Mantsch H.H. (1996). In situ characterization of beta-amyloid in Alzheimer's diseased tissue by synchrotron Fourier transform infrared microspectroscopy. *Biophysical Journal*, 71, 1672 – 1679, ISSN: 0006-3495.
- Diomedea, L., Cassata, G., Fiordaliso, F., Salio, M., Ami, D., Natalello, A., Doglia, S.M., De Luigi, A., Salmona, M. (2010). Tetracycline and its analogues protect *Caenorhabditis elegans* from  $\beta$  amyloid-induced toxicity by targeting oligomers. *Neurobiology of Disease*, 40, 424-431, ISSN: 0969-9961.
- Doglia, S.M., Ami, D., Natalello, A., Gatti-Lafranchi, P., Lotti, M. (2008). Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant proteins within inclusion bodies. *Biotechnology Journal*, 3, 193–201, ISSN: 1860-6768.
- Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikstrom, C., Wold, S. (2006). *Multivariate and Megavariate Data Analysis Basic Principles and Applications*, Umetrics Academy, ISBN-10: 91-973730-2-8, San Jose, CA, USA.
- Evans, M.J. & Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, 154–156, ISSN: 0028-0836.
- Fearn, T. (2002). Discriminant analysis, In: *Handbook of Vibrational Spectroscopy*, Chalmers, J.M. & Griffiths, P.R. (eds.), 2086–2093, ISBN: 0-471-98847-2, Wiley, New York, USA.
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*, Academic Press Professional Inc., ISBN: 0-12-269851-7, San Diego, CA, USA.
- German, M.J., Pollock, H.M., Zhao, B., Tobin, M.J., Hammiche, A., Bentley, A., Cooper, L.J., Martin, F.L., Fullwood, N.J. (2006). Characterization of putative stem cell



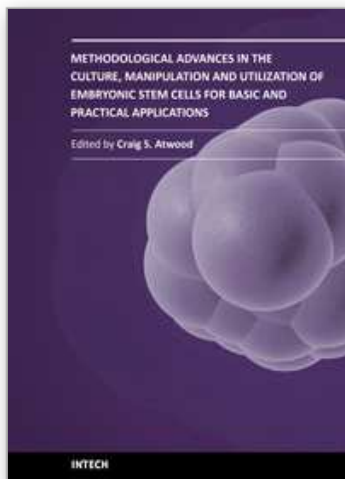
- populations in the cornea using synchrotron infrared microspectroscopy. *Investigative Ophthalmology & Visual Science.*, 47, 2417–2421, ISSN: 0146-0404.
- Hamm, P. (2009). For structural biology, try infrared instead. *Structure*, 17, 149-150, ISSN: 0969-2126.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*, Prentice Hall Inc., ISBN: 0132733501, Englewood Cliffs, New Jersey, USA.
- Héberger, K.; Csomós, E. & Livia S. (2003). Principal Component and Linear Discriminant Analyses of Free Amino Acids and Biogenic Amines in Hungarian Wines. *Journal of agricultural and food chemistry*, 51, 27, 8055-8060, ISSN: 0021-8561.
- Heraud, P. & Tobin, M.J. (2009). The emergence of biospectroscopy in stem cell research. *Stem cell research*, 3, 12-14, ISSN: 1873-5061.
- Heraud, P., Ng, E.S., Caine, S., Yua, Q.C., Hirst, C., Mayberry, R., Bruce, A., Wood, B.R., McNaughton, D., Stanley, E.G., Elefanty, A.G. (2010). Fourier transform infrared microspectroscopy identifies early lineage commitment in differentiating human embryonic stem cells. *Stem Cell Research*, 4, 140-147, ISSN: 1873-5061.
- Jonathan, P., McCarthy, W.V. & Roberts, M.I.A. (1996). Discriminant Analysis With Singular Covariance Matrices. A Method Incorporating Cross-Validation And Efficient Randomized Permutation Tests. *Journal Of Chemometrics*, 10, 4, 189-213, ISSN: 0886-9383.
- Kacurakova, M. & Wilson, R.H. (2001). Developments in mid-infrared FT-IR spectroscopy of selected carbohydrates. *Carbohydrate Polymers*, 44, 291–303, ISSN: 0144-8617.
- Kauppinen, J. K., Moffatt, D.J., Mantsch, H.H., Cameron, D.G. (1981). Fourier self-deconvolution—a method for resolving intrinsically overlapped bands. *Applied Spectroscopy*, 35, 271–276, ISSN: 0003-7028.
- Kazarian, S.G. & Chan, K.L.A. (2006). Applications of ATR-FTIR spectroscopic imaging to biomedical samples. *Biochimica et Biophysica Acta*, 1758, 858–867, ISSN: 0005-2736.
- Konorov, S.O., Schulze, H.G., Caron, N.J., Piret, J.M., Blades, M.W., Turner, R.F.B. (2010). Raman microspectroscopic evidence that dry-fixing preserves the temporal pattern of non-specific differentiation in live human embryonic stem cells. *Journal of Raman spectroscopy*, in press DOI 10.1002/jrs.2769, ISSN: 0377-0486.
- Krafft, C., Steiner, G., Beleites, C., Salzer, R. (2009). Disease recognition by infrared and Raman spectroscopy. *Journal of Biophotonics*, 2, 13-28, ISSN: 1864-063X.
- Krogh, A. (2008). What are artificial neural networks? *Nature Biotechnology*, 25, 2, 195-197, ISSN: 1087-0156.
- Levin, W. & Bhargava, R. (2005). Fourier transform infrared vibrational spectroscopic imaging: Integrating Microscopy and Molecular Recognition. *Annual Review of Physical Chemistry*, 56, 429-474, ISSN: 0066-426X.
- Lewis, E.N., Treado, P.J., Reedor, R.C., Stay, G.M., Dowrey, A.E., Marcott, C., Levint, I.W. (1995). Fourier Transform Spectroscopic Imaging Using an Infrared Focal-Plane Array Detector. *Analytical Chemistry*, 67, 3377-3381, ISSN: 0003-2700.
- Manly, B.F.J. (2004). *Multivariate Statistical Methods*, Chapman & Hall/CRC press, ISBN: 9781584884149, London, UK.
- Miller, L.M. & Dumas P. (2010). From structure to cellular mechanism with infrared microspectroscopy. *Current opinion in structural biology*, 20, 1-8, ISSN: 0959-440X.
- Nakamura, T., Kelly, J.G., Trevisan, J., Cooper, L.J., Bentley, A.J., Carmichael, P.L., Scott, A.D., Cotte, M., Susini, J., Martin-Hirsch, P.L., Kinoshita, S., Fullwood, N.J.,

- Martin, F.L. (2010). Microspectroscopy of spectral biomarkers associated with human corneal stem cells. *Molecular Vision*, 16, 359-368, ISSN: 1090-0535.
- Natalello, A., Ami, D., Brocca, S., Lotti, M., Doglia, S.M. (2005). Secondary structure, conformational stability and glycosylation of a recombinant *Candida rugosa* lipase studied by Fourier-transform infrared spectroscopy. *Biochemical Journal*, 385, 511-517, ISSN: 0264-6021.
- Notingher, I., Bisson, I., Polak, J.M., Hench, L.L. (2004a). In situ spectroscopic study of nucleic acids in differentiating embryonic stem cells. *Vibrational Spectroscopy*, 35, 199-203, ISSN: 0924-2031.
- Notingher, I., Bisson, I., Bishop, A.E., Randle, W.L., Polak, J.M., Hench, L.L. (2004b). In situ spectral monitoring of mRNA translation in embryonic stem cells during differentiation in vitro. *Analytical chemistry*, 76, 3185-3193, ISSN: 0003-2700.
- O'Connell, D., Bertram, J., Giffin, M., Ebesu, J., Empey, C., Mulrooney, M., Wall, M. (2005). Detection of acetylated and non-acetylated histones in chromatin. *Microscopy And Microanalysis*, 11, 1004-1005, ISSN: 1431-9276.
- Orsini, F., Ami, D., Villa, A.M., Sala, G., Bellotti, M.G., Doglia, S.M. (2000). FT-IR microspectroscopy for microbiological studies. *Journal of microbiological methods*, 42, 17-27, ISSN: 0167-7012.
- Oyamada, Y., Komatsu, K., Kimura, H., Mori, M., Oyamada, M. (1996). Differential regulation of gap junction protein (connexin) genes during cardiomyocytic differentiation of mouse embryonic stem cells in vitro. *Experimental Cell Research*, 229, 318-326, ISSN: 0014-4827.
- Pasumarthi, K.B.S. & Field, L.J. (2002). Cardiomyocyte enrichment in differentiating ES cell cultures: strategies and applications, In: *Embryonic Stem Cells Methods and Protocols*, Turksen, K. (Ed.), 157-168, Humana Press, ISBN: 978-0-89603-881-3, New Jersey, USA.
- Pereira, R.C.C., Skrobot, V.L., Castro, E.V.R., Fortes, I.C.P., Pasa, V.M.D. (2006). Determination of Gasoline Adulteration by Principal Components Analysis-Linear Discriminant Analysis Applied to FTIR Spectra. *Energy & Fuels*, 20, 1097-1102, ISSN: 0887-0624.
- Peterson, K.J. & Davidson, E.H. (2000). Regulatory evolution and the origin of the bilaterians. *Proceedings of the National Academy of Sciences*, 97, 4430-4433, ISSN: 0027-8424.
- Rencher, A.C. (2002). *Methods of Multivariate Analysis*, Wiley, ISBN: 978-0-471-41889-4, Hoboken, New Jersey, USA.
- Rezzi, S., Giani, I., Héberger, K., Axelson, D.E., Moretti, V.M., Reniero, F. & Claude Guillou. (2007). Classification of Gilthead Sea Bream (*Sparus aurata*) from <sup>1</sup>H NMR Lipid Profiling Combined with Principal Component and Linear Discriminant Analysis. *Journal of agricultural and food chemistry*, 55, 24, 9963-9968, ISSN: 0021-8561.
- Schulze, H.G., Konorov, S.O., Caron, N.J., Piret, J.M., Blades, M.W., Turner, R.F.B. (2010). Assessing Differentiation Status of Human Embryonic Stem Cells Noninvasively Using Raman Microspectroscopy. *Analytical Chemistry*, 82, 5020-5027, ISSN: 0003-2700.
- Skrobot, L.S., Castro, V.R.E., Pereira, R.C.C., Pasa V.M.D. & Fortes I.C.P. (2007). Use of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) in

- Gas Chromatographic (GC) Data in the Investigation of Gasoline Adulteration. *Energy & Fuels*, 21, 6, 3394-3400, ISSN: 0887-0624.
- Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. *Annual Review of Cell and Developmental Biology*, 17, 435-462, ISSN: 1081-0706.
- Steller, W., Eichenkel, J., Horn, L.C., Braumann, U.D., Binder, H., Salzer, R., Krafft, C. (2006). Delimitation of squamous cell cervical carcinoma using infrared microspectroscopic imaging. *Analytical Bioanalytical Chemistry*, 384, 145-154, ISSN: 1618-2642.
- Susi, H. & Byler, D.M. (1986). Resolution-enhanced Fourier transform infrared spectroscopy of enzymes. *Methods in Enzymology*, 130, 290-311, ISSN: 0076-6879.
- Tamm, L.K. & Tatulian, S.A. (1997). Infrared spectroscopy of proteins and peptides in lipid bilayers. *Quarterly reviews of biophysics*, 30, 365-429, ISSN: 0033-5835
- Tanthanuch, W., Thumanu, K., Lorthongpanich, C., Parnpai, R., Heraud, P. (2010). Neural differentiation of mouse embryonic stem cells studied by FTIR spectroscopy. *Journal of molecular Structure*, 967, 189-195, ISSN: 0022-2860.
- Theophanides, T. & Tajmir-Riahi, H.A. (1985). FT-IR Spectroscopic Evidence of C2'-endo, anti, C3'-endo, anti Sugar Ring Pucker in 5'-GMP and 5'-IMP Nucleotides and their Metal-Adducts, In: *Structure and Motion: Membrane, Nucleic Acids and Proteins*, Clementi, E., Corongiu, G., Sarma, M.H., Sarma R.H. (Eds.), 521-530, Adenine Press, ISBN: 0-940030-12-8, New York, USA.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, 282, 1145-1147, ISSN: 0036-8075.
- Thumanu, K., Tanthanuch, W., Lorthongpanich, C., Heraud, P., Parnpai, R. (2009). FTIR microspectroscopic imaging as a new tool to distinguish chemical composition of mouse blastocyst. *Journal of Molecular Structure*, 933, 104-111, ISSN: 0022-2860.
- Trounson, A. (2006). The Production and Directed Differentiation of Human Embryonic Stem Cells. *Endocrine Reviews*, 27, 208-219, ISSN: 0163-769X.
- Tsuboi, M. (1961), Infrared spectra and secondary structure of deoxyribonucleic acid. *Progress of Theoretical Physics Supplement*, 17, 99-107, ISSN: 0375-9687.
- Vats, A., Bielby, R.C., Tolley, N.S., Nerem, R., Polak, J.M. (2005). Stem cells. *Lancet*, 366, 592-602, ISSN: 0140-6736.
- Verfaillie, C.M. (2002). Adult stem cells: assessing the case for pluripotency. *TRENDS in Cell Biology*, 12, 502-508, ISSN: 0962-8924.
- Walsh, M.J., Hammiche, A., Fellous, T.G., Nicholson, J.M., Cotte, M., Susini, J., Fullwood, N.J., Martin-Hirsch, P.L., Alison, M.R., Martin, F.L. (2009). Tracking the cell hierarchy in the human intestine using biochemical signatures derived by mid-infrared microspectroscopy. *Stem Cell Research*, 3, 15-27, ISSN: 1873-5061.
- Walsh, M.J., Singh, M.N., Pollock, H.M., Cooper, L.J., German, M.J., Stringfellow, H.F., Fullwood, N.J., Paraskeva, E., Martin-Hirsch, P.L., Martin, F.L. (2007). ATR microspectroscopy with multivariate analysis segregates grades of exfoliative cervical cytology. *Biochemical and Biophysical Research Communications*, 352, 213-219, ISSN: 0006-291X.
- Wang, T.D., Triadafilopoulos, G., Crawford, J.M., Dixon, L.R., Bhandari, T., Sahbaie, P., Friedland, S., Soetikno, R., Contag, C.H. (2007). Detection of endogenous

- biomolecules in Barrett's esophagus by Fourier transform infrared spectroscopy. *Proceedings of the National Academy of Sciences*, 104, 15864-15869, ISSN: 0027-8424.
- Wood, B.R., Chernenko, T., Matthäus, C., Diem, M., Chong, C., Bernhard, U., Jene, C., Brandli, A.A., McNaughton, D., Tobin, M.J., Trounson, A., Lacham-Kaplan, O. (2008). Shedding new light on the molecular architecture of oocytes using a combination of synchrotron Fourier transform-infrared and Raman spectroscopic mapping. *Analytical Chemistry*, 80, 9065-9072, ISSN: 0003-2700.
- Zhao, R., Quaroni, L., Casson, A.G. (2010). Fourier transform infrared spectromicroscopic characterization of stem like cell populations in human esophageal normal and adenocarcinoma cell lines. *Analyst*, 135, 53-61, ISSN: 0003-2654.
- Zhizhina, G.P. & Oleinik, E.F. (1972). Infrared spectroscopy of nucleic acids. *Russian Chemical Reviews*, 41, 258 -280, ISSN: 0036-021X.

IntechOpen



## **Methodological Advances in the Culture, Manipulation and Utilization of Embryonic Stem Cells for Basic and Practical Applications**

Edited by Prof. Craig Atwood

ISBN 978-953-307-197-8

Hard cover, 506 pages

**Publisher** InTech

**Published online** 26, April, 2011

**Published in print edition** April, 2011

Pluripotent stem cells have the potential to revolutionise medicine, providing treatment options for a wide range of diseases and conditions that currently lack therapies or cures. This book describes methodological advances in the culture and manipulation of embryonic stem cells that will serve to bring this promise to practice.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Diletta Ami, Paolo Mereghetti, Antonino Natalello and Silvia M. Doglia (2011). Fourier Transform Infrared Microspectroscopy as a Tool for Embryonic Stem Cell Studies, Methodological Advances in the Culture, Manipulation and Utilization of Embryonic Stem Cells for Basic and Practical Applications, Prof. Craig Atwood (Ed.), ISBN: 978-953-307-197-8, InTech, Available from: <http://www.intechopen.com/books/methodological-advances-in-the-culture-manipulation-and-utilization-of-embryonic-stem-cells-for-basic-and-practical-applications/fourier-transform-infrared-microspectroscopy-as-a-tool-for-embryonic-stem-cell-studies>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821



© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen